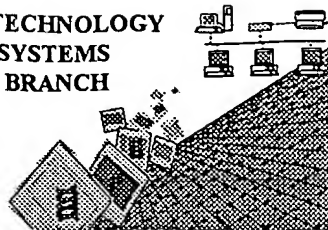


BIOTECHNOLOGY
SYSTEMS
BRANCH



0590
1202

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/081,555B
Source: O/PK
Date Processed by STIC: 12/13/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER** **VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efb/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002



OIPE

RAW SEQUENCE LISTING

DATE: 12/13/2002

PATENT APPLICATION: US/10/081,555B

TIME: 14:59:34

Input Set : A:\Salk2270.app

Output Set: N:\CRF4\12132002\J081555B.raw

3 <110> APPLICANT: EVANS, RONALD M.
 5 <120> TITLE OF INVENTION: NOVEL STEROID-ACTIVATED NUCLEAR RECEPTORS AND USES THEREFOR
 7 <130> FILE REFERENCE: SALK2270-5
 9 <140> CURRENT APPLICATION NUMBER: 10/081,555B
 10 <141> CURRENT FILING DATE: 2002-02-20
 12 <150> PRIOR APPLICATION NUMBER: 09/458,366
 13 <151> PRIOR FILING DATE: 1999-12-09
 15 <150> PRIOR APPLICATION NUMBER: 09/227,718
 16 <151> PRIOR FILING DATE: 1999-01-08
 18 <150> PRIOR APPLICATION NUMBER: 09/005,286
 19 <151> PRIOR FILING DATE: 1998-01-09
 21 <160> NUMBER OF SEQ ID NOS: 48
 23 <170> SOFTWARE: PatentIn Ver. 2.1
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 2068
 27 <212> TYPE: DNA
 28 <213> ORGANISM: Homo sapiens
 30 <220> FEATURE:
 31 <221> NAME/KEY: CDS
 32 <222> LOCATION: (583)..(1884)
 34 <220> FEATURE:
 35 <221> NAME/KEY: modified_base
 36 <222> LOCATION: (1263)
 37 <223> OTHER INFORMATION: a, c, t, or g
 39 <400> SEQUENCE: 1
 40 ggcacgagga gatctaggtt caaattaatg ttgcccttag ttgtaaagga cagagaccct 60
 42 cagactgatg aaatgcgctc agaattactt agacaaagcg gatatttgcc actctcttcc 120
 44 ccttttcctg tgtttttgta gtgaagagac ctgaaagaaa aaagtaggga gaacataatg 180
 46 agaacaataa cggtaatctc ttcatattgct agttcaagtg ctggacttg gacttaggag 240
 48 gggcaatgga gccgcttagt gcctacatct gacttggaact gaaatatagg tgagagacaa 300
 50 gattgtctca tatccgggga aatcataacc tatgactagg acgggaagag gaagcactgc 360
 52 ctttacttca gtgggaatct cggcctcagc ctgcaagcca agtggttcaca gtgagaaaag 420
 54 caagagaata agctaatact cctgtcctga acaaggcagc ggctccttg taaagctact 480
 56 ccttgatcga tcctttgcac cggattgttc aaagtggacc ccaggggaga agtcggagca 540
 58 aagaacttac caccaagcag tccaagaggc ccagaagcaa ac ctg gag gtg aga 594
 59 Met Glu Val Arg
 60 1
 62 ccc aaa gaa agc tgg aac cat gct gac ttt gta cac tgt gag gac aca 642
 63 Pro Lys Glu Ser Trp Asn His Ala Asp Phe Val His Cys Glu Asp Thr
 64 5 10 15 20
 66 gag tct gtt cct gga aag ccc agt gtc aac gca gat gag gaa gtc gga 690
 67 Glu Ser Val Pro Gly Lys Pro Ser Val Asn Ala Asp Glu Glu Val Gly
 68 25 30 35

see pp 1-4, 6

**Does Not Comply
Corrected Diskette Needed**

**Does Not Comply
Corrected Diskette Needed**

see p.2 "c" is at location 1263 "c" can

*only
represent
itself,
nothing
else*

*use "n"
instead
and explain*

*in 2207-2237
section*

RAW SEQUENCE LISTING

DATE: 12/13/2002

PATENT APPLICATION: US/10/081,555B

TIME: 14:59:35

Input Set : A:\Salk2270.app

Output Set: N:\CRF4\12132002\J081555B.raw

```

70 ggt ccc caa atc tgc cgt gta tgt ggg gac aag gcc act ggc tat cac 738
71 Gly Pro Gln Ile Cys Arg Val Cys Gly Asp Lys Ala Thr Gly Tyr His
72          40          45          50
74 ttc aat gtc atg aca tgt gaa gga tgc aag ggc ttt ttc agg agg gcc 786
75 Phe Asn Val Met Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ala
76          55          60          65
78 atg aaa cgc aac gcc cgg ctg agg tgc ccc ttc cgg aag ggc gcc tgc 834
79 Met Lys Arg Asn Ala Arg Leu Arg Cys Pro Phe Arg Lys Gly Ala Cys
80          70          75          80
82 gag atc acc cgg aag acc cgg cga cag tgc cag gcc tgc cgc ctg cgc 882
83 Glu Ile Thr Arg Lys Thr Arg Arg Gln Cys Gln Ala Cys Arg Leu Arg
84 85          90          95          100
86 aag tgc ctg gag agc ggc atg aag aag gag atg atc atg tcc gac gag 930
87 Lys Cys Leu Glu Ser Gly Met Lys Lys Glu Met Ile Met Ser Asp Glu
88          105          110          115
90 gcc gtg gag gag agg cgg gcc ttg atc aag cgg aag aaa agt gaa cgg 978
91 Ala Val Glu Glu Arg Arg Ala Leu Ile Lys Arg Lys Lys Ser Glu Arg
92          120          125          130
94 aca ggg act cag cca ctg gga gtg cag ggg ctg aca gag gag cag cgg 1026
95 Thr Gly Thr Gln Pro Leu Gly Val Gln Gly Leu Thr Glu Glu Gln Arg
96          135          140          145
98 atg atg atc agg gag ctg atg gac gct cag atg aaa acc ttt gac act 1074
99 Met Met Ile Arg Glu Leu Met Asp Ala Gln Met Lys Thr Phe Asp Thr
100          150          155          160
102 acc ttc tcc cat ttc aag aat ttc cgg ctg cca ggg gtg ctt agc agt 1122
103 Thr Phe Ser His Phe Lys Asn Phe Arg Leu Pro Gly Val Leu Ser Ser
104 165          170          175          180
106 ggc tgc gag ttg cca gag tct ctg cag gcc cca tgc agg gaa gaa gct 1170
107 Gly Cys Glu Leu Pro Glu Ser Leu Gln Ala Pro Ser Arg Glu Glu Ala
108          185          190          195
110 gcc aag tgg agc cag gtc cgg aaa gat ctg tgc tct ttg aag gtc tct 1218
111 Ala Lys Trp Ser Gln Val Arg Lys Asp Leu Cys Ser Leu Lys Val Ser
112          200          205          210
114 ctg cag ctg cgg ggg gag gat ggc agt gtc tgg aac tac aaa ccc cca 1266
115 Leu Gln Leu Arg Gly Glu Asp Gly Ser Val Trp Asn Tyr Lys Pro Pro
116          215          220          225
118 gcc gac agt ggc ggg aaa gag atc ttc tcc ctg ctg ccc cac atg gct 1314
119 Ala Asp Ser Gly Gly Lys Glu Ile Phe Ser Leu Leu Pro His Met Ala
120          230          235          240
122 gac atg tca acc tac atg ttc aaa ggc atc atc agc ttt gcc aaa gtc 1362
123 Asp Met Ser Thr Tyr Met Phe Lys Gly Ile Ile Ser Phe Ala Lys Val
124 245          250          255          260
126 atc tcc tac ttc agg gac ttg ccc atc gag gac cag atc tcc ctg ctg 1410
127 Ile Ser Tyr Phe Arg Asp Leu Pro Ile Glu Asp Gln Ile Ser Leu Leu
128          265          270          275
130 aag ggg gcc gct ttc gag ctg tgt caa ctg aga ttc aac aca gtg ttc 1458
131 Lys Gly Ala Ala Phe Glu Leu Cys Gln Leu Arg Phe Asn Thr Val Phe
132          280          285          290
134 aac gcg gag act gga acc tgg gag tgt ggc cgg ctg tcc tac tgc ttg 1506

```

RAW SEQUENCE LISTING

DATE: 12/13/2002

PATENT APPLICATION: US/10/081,555B

TIME: 14:59:35

Input Set : A:\Salk2270.app

Output Set: N:\CRF4\12132002\J081555B.raw

```

135 Asn Ala Glu Thr Gly Thr Trp Glu Cys Gly Arg Leu Ser Tyr Cys Leu
136          295          300          305
138 gaa gac act gca ggt ggc ttc cag caa ctt cta ctg gag ccc atg ctg 1554
139 Glu Asp Thr Ala Gly Gly Phe Gln Gln Leu Leu Leu Glu Pro Met Leu
140      310          315          320
142 aaa ttc cac tac atg ctg aag aag ctg cag ctg cat gag gag gag tat 1602
143 Lys Phe His Tyr Met Leu Lys Lys Leu Gln Leu His Glu Glu Glu Tyr
144 325          330          335          340
146 gtg ctg atg cag gcc atc tcc ctc ttc tcc cca gac cgc cca ggt gtg 1650
147 Val Leu Met Gln Ala Ile Ser Leu Phe Ser Pro Asp Arg Pro Gly Val
148          345          350          355
150 ctg cag cac cgc gtg gtg gac cag ctg cag gag caa ttc gcc att act 1698
151 Leu Gln His Arg Val Val Asp Gln Leu Gln Glu Gln Phe Ala Ile Thr
152          360          365          370
154 ctg aag tcc tac att gaa tgc aat cgg ccc cag cct gct cat agg ttc 1746
155 Leu Lys Ser Tyr Ile Glu Cys Asn Arg Pro Gln Pro Ala His Arg Phe
156          375          380          385
158 ttg ttc ctg aag atc atg gct atg ctc acc gag ctc cgc agc atc aat 1794
159 Leu Phe Leu Lys Ile Met Ala Met Leu Thr Glu Leu Arg Ser Ile Asn
160      390          395          400
162 gct cag cac acc cag cgg ctg ctg cgc atc cag gac ata caa ccc ttt 1842
163 Ala Gln His Thr Gln Arg Leu Leu Arg Ile Gln Asp Ile His Pro Phe
164 405          410          415          420
166 gct acg ccc ctc atg cag gag ttg ttc ggt atc aca ggt agc tga 1887
167 Ala Thr Pro Leu Met Gln Glu Leu Phe Gly Ile Thr Gly Ser
168          425          430
170 gtggctgtcc ttgggtgaca cctccgagag gtatgttagac ccagagccct ctgagtcgcc 1947
172 actcccgggc caagacagat ggacactgcc aagagccgac aatgccctgc tggcctgtct 2007
174 ccctagggaa ttcctgctat gacagctggc tagcattcct caggaaggac atgggggtgcc 2067
176 c 2068
179 <210> SEQ ID NO: 2
180 <211> LENGTH: 434
181 <212> TYPE: PRT
182 <213> ORGANISM: Homo sapiens
184 <220> FEATURE:
185 <221> NAME/KEY: MOD_RES
186 <222> LOCATION: (227)
187 <223> OTHER INFORMATION: Threonine
189 <400> SEQUENCE: 2
190 Met Glu Val Arg Pro Lys Glu Ser Trp Asn His Ala Asp Phe Val His
191 1 5 10 15
193 Cys Glu Asp Thr Glu Ser Val Pro Gly Lys Pro Ser Val Asn Ala Asp
194 20 25 30
196 Glu Glu Val Gly Gly Pro Gln Ile Cys Arg Val Cys Gly Asp Lys Ala
197 35 40 45
199 Thr Gly Tyr His Phe Asn Val Met Thr Cys Glu Gly Cys Lys Gly Phe
200 50 55 60
202 Phe Arg Arg Ala Met Lys Arg Asn Ala Arg Leu Arg Cys Pro Phe Arg
203 65 70 75 80

```

Pro is at location 227, and can only represent itself

P.4

Use "Xaa" instead and explain in (2207-2223) section

IF the

amino acid at location

227 can be either Pro or thr

RAW SEQUENCE LISTING

DATE: 12/13/2002

PATENT APPLICATION: US/10/081,555B

TIME: 14:59:35

Input Set : A:\Salk2270.app

Output Set: N:\CRF4\12132002\J081555B.raw

```

205 Lys Gly Ala Cys Glu Ile Thr Arg Lys Thr Arg Arg Gln Cys Gln Ala
206      85      90      95
208 Cys Arg Leu Arg Lys Cys Leu Glu Ser Gly Met Lys Lys Glu Met Ile
209      100      105      110
211 Met Ser Asp Glu Ala Val Glu Glu Arg Arg Ala Leu Ile Lys Arg Lys
212      115      120      125
214 Lys Ser Glu Arg Thr Gly Thr Gln Pro Leu Gly Val Gln Gly Leu Thr
215      130      135      140
217 Glu Glu Gln Arg Met Met Ile Arg Glu Leu Met Asp Ala Gln Met Lys
218 145      150      155      160
220 Thr Phe Asp Thr Thr Phe Ser His Phe Lys Asn Phe Arg Leu Pro Gly
221      165      170      175
223 Val Leu Ser Ser Gly Cys Glu Leu Pro Glu Ser Leu Gln Ala Pro Ser
224      180      185      190
226 Arg Glu Glu Ala Ala Lys Trp Ser Gln Val Arg Lys Asp Leu Cys Ser
227      195      200      205
229 Leu Lys Val Ser Leu Gln Leu Arg Gly Glu Asp Gly Ser Val Trp Asn
230      210      215      220
232 Tyr Lys Pro Pro Ala Asp Ser Gly Gly Lys Glu Ile Phe Ser Leu Leu
233 225      230      235      240
235 Pro His Met Ala Asp Met Ser Thr Tyr Met Phe Lys Gly Ile Ile Ser
236      245      250      255
238 Phe Ala Lys Val Ile Ser Tyr Phe Arg Asp Leu Pro Ile Glu Asp Gln
239      260      265      270
241 Ile Ser Leu Leu Lys Gly Ala Ala Phe Glu Leu Cys Gln Leu Arg Phe
242      275      280      285
244 Asn Thr Val Phe Asn Ala Glu Thr Gly Thr Trp Glu Cys Gly Arg Leu
245      290      295      300
247 Ser Tyr Cys Leu Glu Asp Thr Ala Gly Gly Phe Gln Gln Leu Leu Leu
248 305      310      315      320
250 Glu Pro Met Leu Lys Phe His Tyr Met Leu Lys Lys Leu Gln Leu His
251      325      330      335
253 Glu Glu Glu Tyr Val Leu Met Gln Ala Ile Ser Leu Phe Ser Pro Asp
254      340      345      350
256 Arg Pro Gly Val Leu Gln His Arg Val Val Asp Gln Leu Gln Glu Gln
257      355      360      365
259 Phe Ala Ile Thr Leu Lys Ser Tyr Ile Glu Cys Asn Arg Pro Gln Pro
260      370      375      380
262 Ala His Arg Phe Leu Phe Leu Lys Ile Met Ala Met Leu Thr Glu Leu
263 385      390      395      400
265 Arg Ser Ile Asn Ala Gln His Thr Gln Arg Leu Leu Arg Ile Gln Asp
266      405      410      415
268 Ile His Pro Phe Ala Thr Pro Leu Met Gln Glu Leu Phe Gly Ile Thr
269      420      425      430
271 Gly Ser
274 <210> SEQ ID NO: 3
275 <211> LENGTH: 25
276 <212> TYPE: DNA
277 <213> ORGANISM: Artificial Sequence

```

RAW SEQUENCE LISTING

DATE: 12/13/2002

PATENT APPLICATION: US/10/081,555B

TIME: 14:59:35

Input Set : A:\Salk2270.app

Output Set: N:\CRF4\12132002\J081555B.raw

```

279 <220> FEATURE:
280 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
281     response element from the steroid hydroxylase,
282     rCYP3A1
284 <400> SEQUENCE: 3
285 tagacagttc atgaagttca tctac 25
288 <210> SEQ ID NO: 4
289 <211> LENGTH: 25
290 <212> TYPE: DNA
291 <213> ORGANISM: Artificial Sequence
293 <220> FEATURE:
294 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
295     response element from the steroid hydroxylase,
296     rCYP3A2
298 <400> SEQUENCE: 4
299 taagcagttc ataaagttca tctac 25
302 <210> SEQ ID NO: 5
303 <211> LENGTH: 25
304 <212> TYPE: DNA
305 <213> ORGANISM: Artificial Sequence
307 <220> FEATURE:
308 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
309     response element from the steroid hydroxylase,
310     rUGT1A6
312 <400> SEQUENCE: 5
313 actgtagttc ataaagttca catgg 25
316 <210> SEQ ID NO: 6
317 <211> LENGTH: 26
318 <212> TYPE: DNA
319 <213> ORGANISM: Artificial Sequence
321 <220> FEATURE:
322 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
323     response element from the steroid hydroxylase,
324     rbCYP2C1
326 <400> SEQUENCE: 6
327 caatcagttc aacagggttc accaat 26
330 <210> SEQ ID NO: 7
331 <211> LENGTH: 33
332 <212> TYPE: DNA
333 <213> ORGANISM: Artificial Sequence
335 <220> FEATURE:
336 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
337     response element from the steroid hydroxylase,
338     rP450R
340 <400> SEQUENCE: 7
341 cacaggtgag ctgaggccag cagcaggtcg aaa 33
344 <210> SEQ ID NO: 8
345 <211> LENGTH: 27
346 <212> TYPE: DNA

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/081,555B

DATE: 12/13/2002
TIME: 14:59:36

Input Set : A:\Salk2270.app
Output Set: N:\CRF4\12132002\J081555B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:22; N Pos. 7,8,9,10,11
Seq#:23; N Pos. 7,8,9,10,11,12
Seq#:44; N Pos. 7,8,9
Seq#:45; N Pos. 7,8,9,10
Seq#:46; N Pos. 7,8,9,10,11
Seq#:47; N Pos. 7,8,9
Seq#:48; N Pos. 7,8,9,10

VERIFICATION SUMMARY

DATE: 12/13/2002

PATENT APPLICATION: US/10/081,555B

TIME: 14:59:36

Input Set : A:\Salk2270.app

Output Set: N:\CRF4\12132002\J081555B.raw

L:548 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0
L:567 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:0
L:840 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44 after pos.:0
L:858 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:45 after pos.:0
L:876 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46 after pos.:0
L:894 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47 after pos.:0
L:912 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48 after pos.:0